

## **REMARKS**

Claims 1-16 are now pending in the application. Claims 1, 8, and 15 are amended. The Examiner is respectfully requested to reconsider and withdraw the rejections in view of the amendments and remarks contained herein. Specifically, Applicant asserts the following: (1) that Laurila et al. only teaches trying to recognize a model trained on noise and non-command words as an estimate of noise against which a command word recognition score is compared; (2) that Chan only teaches searching a range of sound signal energy values for a minimum value, and not a range of recognition scores; (3) that the differences between Applicant's claimed invention and the teachings of the references relied upon by the Examiner are significant because Applicant's claimed invention can dynamically adjust for changes in background environment by forcing the matching of the word model with the background environment when the word is not spoken in temporal proximity to speaking of the word (i.e., just before and just after the word is spoken); (4) that an essential component to the process of Applicant's claimed invention is estimation of the background score with reference to which confidence in presence of a word is determined based on the recognition score for the word; (5) that all of the independent claims recite limitations to the aforementioned essential component; (6) that none of the cited references, alone or combined, teach, suggest, or motivate this process or provide this capability; (7) that support for the amendments and arguments detailed herein may be found in the specification as originally filed at paragraph [0021]; and (8) that amendments to the claims are not narrowing amendments because they explicitly recite subject matter already inherent in the amended claims as originally filed given any reasonable

interpretation of the claim language in view of the teachings of the specification as originally filed.

**REJECTION UNDER 35 U.S.C. § 112**

Claims 1-16 stand rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point and distinctly claim the subject matter which Applicant regards as the invention. This rejection is respectfully traversed.

The Examiner remarks that claims 1, 2, 12, and 15 incorrectly recite “absolute likelihood” when they should recite “absolute value of log likelihood”. However, claims 2 and 12 do not recite “absolute likelihood”. Thus, claim set 12-14 should not have been rejected on these grounds, since the term “absolute likelihood” does not appear in independent claim 12 or anywhere in the chain of dependency of this claim set. However, the term “absolute likelihood” does appear in claims 1 and 15, as well as in claim 8. Applicant has amended these claims as suggested by the Examiner, and further asserts that these amendments are not narrowing amendments because they explicitly recite subject matter already inherent in the amended claims as originally filed given any reasonable interpretation of the claim language in view of the teachings of the specification as originally filed.

Accordingly, Applicant believes the rejection of claims 1-11 and 15-16 has been rendered moot. Accordingly, Applicant requests the Examiner withdraw the rejection of claims 1-22 on these grounds.

## **REJECTION UNDER 35 U.S.C. § 102**

Claims 1,2,6,7,8,9,10,11, and 15 stand rejected under 35 U.S.C. § 102(b) as being anticipated by Laurila et al. (EP 1 020 847 A2). This rejection is respectfully traversed.

Laurila et al. is generally directed toward multistage speech recognition using confidence measures. In particular, Laurila et al. is directed toward a non-continuous command word recognition technique that begins by attempting to recognize a command word within a predetermined amount of time, and conditionally extends the amount of time to attempt to recognize a repetition of the command word based on plural confidence thresholds (Abstract). If confidence is especially high or low that the spoken word is a command word (i.e., above a high confidence or below a low confidence threshold), then the word is recognized or not recognized and the recognition phase is exited. However, if the confidence is between the two thresholds, then time is extended to allow the user to repeat the command word. The confidence score is obtained by comparing (via subtraction) a highest recognition score obtained by attempting to recognize models trained on command words, to a recognition score obtained by attempting to recognize a model trained on noise and non-command words (column 6, lines 44-57). However, Laurila et al. does not teach calculating a first confidence score tracking a noise-corrected likelihood that a first word is in a speech signal based on a matching ratio between: (1) a first minimum recognition value of a first recognition score tracking an absolute value of log likelihood that the first word is in the speech signal; and (2) a first background score estimated based on the first recognition score.

Applicant's claimed invention is generally directed toward robust word spotting. In particular, Applicant's claimed invention estimates background noise based on a recognition score obtained by an attempt to recognize a word, and measures confidence in recognition of the word by comparing an extreme value of the recognition score to the estimate of the background noise. For example, independent claim 1 as amended recites "calculating a first confidence score based on a matching ratio between a first minimum recognition value of the first recognition score and the first background score, the first confidence score tracking a noise-corrected likelihood that the first word is in the speech signal." Other elements of independent claim 1 already specify that the first recognition score tracks an absolute value of log likelihood that the first word is in the speech signal, and that the first background score is estimated based on the first recognition score. Accordingly, independent claim 1 as amended essentially recites calculating a first confidence score tracking a noise-corrected likelihood that a first word is in a speech signal based on a matching ratio between: (1) a first minimum recognition value of a first recognition score tracking an absolute value of log likelihood that the first word is in the speech signal; and (2) a first background score estimated based on the first recognition score. Independent claim 15 as amended recites similar limitations.

In rejecting independent claims 1 and 15 as originally filed, the Examiner perhaps assumes that Laurila et al. teaches modeling noise based on attempts to recognize command words. However, this reading would constitute impermissible hindsight reasoning on the Examiner's part, especially since it is at least as reasonable to view Laurila et al. as teaching training a recognition model on noise and non-command

words, and attempting to recognize that model on the speech signal as a measure of background noise. However, perhaps the Examiner reads the claims on Laurila et al. as follows: (1) generating a first recognition score by attempting to recognize a non-command word; (2) estimating a background score based on the first recognition score by treating the first recognition score as the background score; (3) generating a second recognition score by attempting to recognize a command word; (4) calculating a first confidence score based on a matching ratio between a first minimum recognition value of the second recognition score (i.e., the command word) and the first background score (i.e., the non-command word); and (5) deeming that the first confidence score tracks a noise-corrected likelihood that the first word (i.e., the non-command word) is in the speech signal when confidence is low. However, even this reading is not really correct because the non-command word is part of the noise model in Laurila et al.; therefore, the first confidence score cannot track a noise-corrected likelihood that the non-command word is in the speech signal under this reading.

Nevertheless, even if one were to erroneously deem the reading detailed above as correct, the amendments to claims 1 and 15 specify that the confidence score is determined on a matching ratio between a first minimum recognition value of the first recognition score and the first background score that was estimated based on the first recognition score, wherein the first recognition score is obtained by attempting to recognize a first word, and the confidence score evaluates presence of the first word in the speech signal. Thus, the claims as amended cannot be viewed to read on Laurila et al. as suggested by the Examiner. Applicant respectfully asserts that the first minimum value recited in the claims as originally filed inherently pertains to the first recognition

score, and that Applicant could not be expected to reasonably foresee the unorthodox and highly creative reading of the claims as suggested by the Examiner. Accordingly, Applicant respectfully asserts that the amendments to claims 1 and 15 are not narrowing amendments, but explicitly recite subject matter that was already inherent in the amended claims as originally filed given any reasonable interpretation of the claim language in view of the teachings of the specification as originally filed.

If Applicant has misunderstood the Examiner's suggested reading of the claims on Laurila et al., Applicant respectfully requests the Examiner explain the suggested reading in more detail. In particular, Applicant requests more explanation regarding how the background score in Laurilla et al. is estimated based on a recognition score of a word, with a confidence measure that the word is present being obtained by comparing the recognition score to the background score.

Applicant believes that claims 1 and 15 distinguish over the teachings of Laurila et al., especially as amended. Accordingly, Applicant respectfully requests the Examiner withdraw the rejections of independent claims 1 and 15 under 35 U.S.C. § 102(a) on these grounds, along with rejection on these grounds of all claims dependent therefrom.

#### **REJECTION UNDER 35 U.S.C. § 103**

Claims 3, 12, and 16 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Laurila et al. (EP 1 020 847 A2) in view of Modi et al. (U.S. Pat. No. 6,125,345). This rejection is respectfully traversed.

Applicant respectfully refers the Examiner to remarks detailed above with respect to rejection under 35 U.S.C. § 102(a). Applicant respectfully asserts that independent

claim 12 recites at least the limitations recited in independent claims 1 and 15 when claim 12 recites "dividing a minimum value of a speech recognition score by an average value of the speech recognition score over a predetermined period of time such that a matching ratio results, the average value defining an estimated background score". Applicant further notes that the Examiner only relies on Modi et al. to teach normalizing confidence scores. Thus, Modi et al., in combination with Laurilla et al., does not teach the limitations recited in claim 12 or allowable base claims 1 and 15.

The differences between Applicant's claimed invention and the teachings of the references relied upon by the Examiner are significant because Applicant's claimed invention can dynamically adjust for changes in background environment by forcing the matching of the word model with the background environment when the word is not spoken in temporal proximity to speaking of the word (i.e., just before and just after the word is spoken). An essential component to this process is estimation of the background score with reference to which confidence in presence of a word is determined based on the recognition score for the word. None of the cited references, alone or combined, teach, suggest, or motivate this process or provide this capability.

Accordingly, Applicant respectfully requests the Examiner withdraw the rejections of independent claim 12 under 35 U.S.C. § 103(a) because it is in condition for allowance. Further, Applicant respectfully requests the Examiner withdraw the rejections of dependent claims 3 and 16 under 35 U.S.C. § 103(a) based on their dependency from allowable base claims.

Claims 4 and 13 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Laurila et al. (EP 1 020 847 A2) in view of Modi et al. (U.S. Pat. No. 6,125,345) and

further in view of Junkawitsch (U.S. Pat. No. 6,505,156). This rejection is respectfully traversed.

Applicant respectfully refers the Examiner to remarks detailed above with respect to rejection under 35 U.S.C. § 102(a). Applicant also refers the Examiner to remarks detailed above with respect to rejection of independent claim 12 under 35 U.S.C. § 103(a). Applicant Further notes that the Examiner only relies on Modi et al. to teach normalizing confidence scores. Applicant yet further notes that the Examiner only relies on Junkawitsch to teach use of a minimum score to determine the optimal score in continuous speech. Thus, Junkawitsch, in combination with Modi et al., and Laurilla et al., do not teach the limitations recited in allowable base claims 1 and 12. The differences between Applicant's claimed invention and the teachings of the references relied upon by the Examiner are significant because Applicant's claimed invention can dynamically adjust for changes in background environment by forcing the matching of the word model with the background environment when the word is not spoken in temporal proximity to speaking of the word (i.e., just before and just after the word is spoken). An essential component to this process is estimation of the background score with reference to which confidence in presence of a word is determined based on the recognition score for the word. None of the cited references, alone or combined, teach, suggest, or motivate this process or provide this capability.

Accordingly, Applicant respectfully requests the Examiner withdraw the rejections of dependent claims 4 and 13 under 35 U.S.C. § 103(a) based on their dependency from allowable base claims.



Claims 5 and 14 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Laurila et al. (EP 1 020 847 A2) in view of Modi et al. (U.S. Pat. No. 6,125,345), further in view of Junkawitsch (U.S. Pat. No. 6,505,156), and yet further in view of Chan (U.S. Pat. No. 6,032,114). This rejection is respectfully traversed.

Applicant respectfully refers the Examiner to remarks detailed above with respect to rejection under 35 U.S.C. § 102(a). Applicant also refers the Examiner to remarks detailed above with respect to rejection of independent claim 12 under 35 U.S.C. § 103(a). Applicant further notes that the Examiner only relies on Modi et al. to teach normalizing confidence scores. Applicant yet further notes that the Examiner only relies on Junkawitsch to teach use of a minimum score to determine the optimal score in continuous speech. Applicant still further notes that the Examiner only relies on Chan to teach searching a predetermined time frame for a minimum value. However, Applicant respectfully asserts that the Examiner errs in finding that Chan teaches searching a range of recognition scores; Chan teaches searching a range of sound signal energy values. Thus, the suggested combination of Chan, Junkawitsch, Modi et al., and Laurilla et al., does not teach the limitations recited in allowable base claims 1 and 12.

The differences between Applicant's claimed invention and the teachings of the references relied upon by the Examiner are significant because Applicant's claimed invention can dynamically adjust for changes in background environment by forcing the matching of the word model with the background environment when the word is not spoken in temporal proximity to speaking of the word (i.e., just before and just after the word is spoken). An essential component to this process is estimation of the background score with reference to which confidence in presence of a word is

determined based on the recognition score for the word. None of the cited references, alone or combined, teach, suggest, or motivate this process or provide this capability.

Accordingly, Applicant respectfully requests the Examiner withdraw the rejections of dependent claims 5 and 14 under 35 U.S.C. § 103(a) based on their dependency from allowable base claims.

**CONCLUSION**

It is believed that all of the stated grounds of rejection have been properly traversed, accommodated, or rendered moot. Applicant therefore respectfully requests that the Examiner reconsider and withdraw all presently outstanding rejections. It is believed that a full and complete response has been made to the outstanding Office Action, and as such, the present application is in condition for allowance. Thus, prompt and favorable consideration of this amendment is respectfully requested. If the Examiner believes that personal communication will expedite prosecution of this application, the Examiner is invited to telephone the undersigned at (248) 641-1600.

Respectfully submitted,

Dated: Aug 5, 2004

By: Greg Stobbs  
Gregory Stobbs  
Reg. No. 28,764

HARNES, DICKEY & PIERCE, P.L.C.  
P.O. Box 828  
Bloomfield Hills, Michigan 48303  
(248) 641-1600